

Reducción del Riesgo en Muestras Intencionales

Mariana Tafur Arciniegas, Septiembre 5 de 2016

Antecedentes

➤ Motivación:

Pros y contras tanto de muestras grandes como pequeñas

➤ Manejo del sesgo en la selección de muestras

➤ Probabilidad vs. Propósito (Sandelowski, 1995)

Investigadores cualitativos valoran una comprensión profunda que da la información de casos particulares, mientras los investigadores cuantitativos valoran poder generalizar a grandes poblaciones

Objetivos de Investigación

- Reducir parcialidad en el muestreo intencional
 - Este muestreo puede realizarse utilizando análisis estadístico
 - Datos atípicos representan casos críticos en una muestra de máxima variación
 - El tamaño de la muestra afecta la calidad de la identificación de casos críticos

Marco Teórico

- Manejo de “Outliers” o datos atípicos
 - Casos Críticos como metodología de investigación (ej. Gremler, 2004)
 - Manejo de ruido (ej. Cherednichenko, 2005)

- Cálculo para identificación de datos atípicos
 - Basado en la distancia, controlando variabilidad (Cherednichenko, 2005)

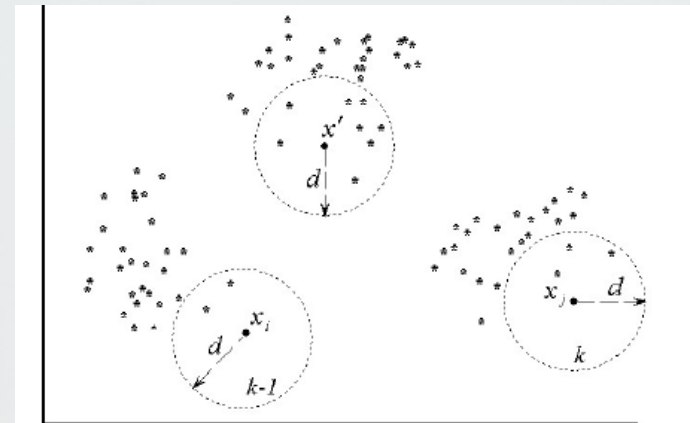


Figure 5. Illustration of outlier definition by Knorr and Ng.

Tomado de Cherednichenko, 2005. Cap.3

Cálculo de Outliers

Limitaciones en nubes basadas en distancia

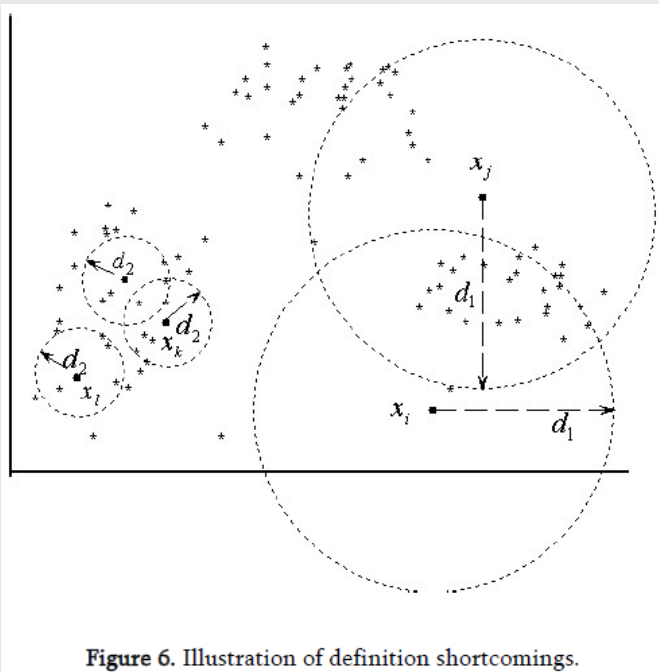


Figure 6. Illustration of definition shortcomings.

Tomado de Cherednichenko, 2005. Cap.3

Inclusión de la Variabilidad

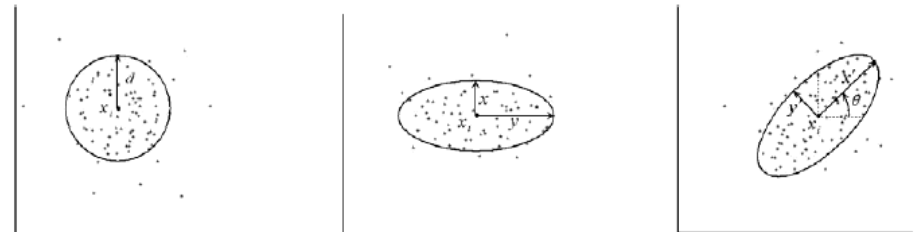


Figure 8. Data objects of the same scale and variability (left). Different scales, variability and no correlation (middle). Different scales, variability and correlation (right).

Tomado de Cherednichenko, 2005. Cap.3

Muestreo Intencional

- Casos extremos
- Muestreo por intensidad
- Máxima variación
- Muestra homogénea
- Casos típicos
- Casos críticos
- Bola de nieve
- Muestra por criterio

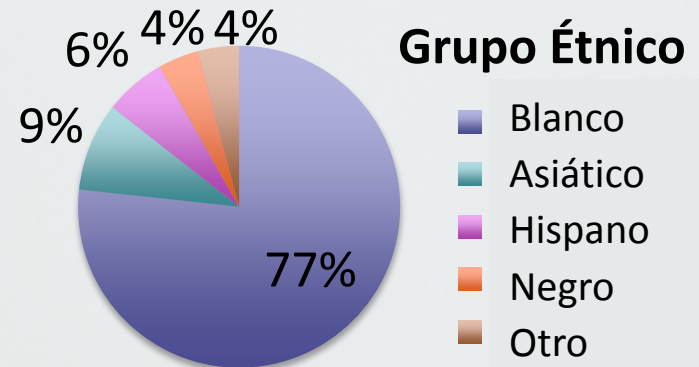
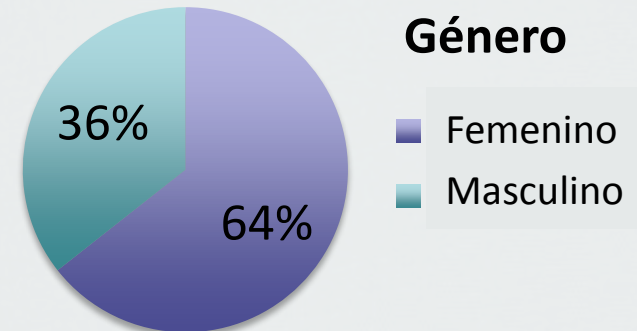
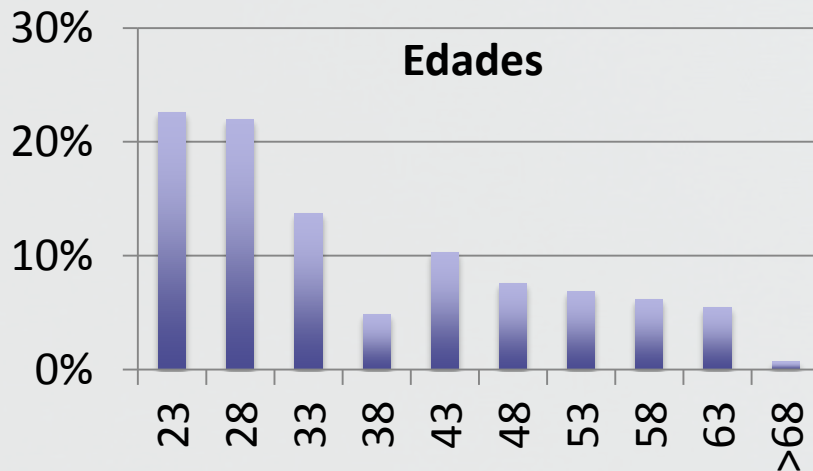
(Suri, 2011)

Método

- **Aproximación pragmática:** Diseño anidado para muestreo cuantitativo en diseños cualitativos (Tashakkori & Teddlie, 2003)
- **Selección de Participantes:** Muestreo por conveniencia en una universidad y una compañía en Estados Unidos central (Midwest)
- **Dos variables de análisis:** Aprendizaje a través de la vida y perfil profesional

Descripción de los Participantes

- Muestra inicial: 146 participantes
- Muestra válida: 118 participantes



Recolección de Datos

➤ **Perfil Profesional**

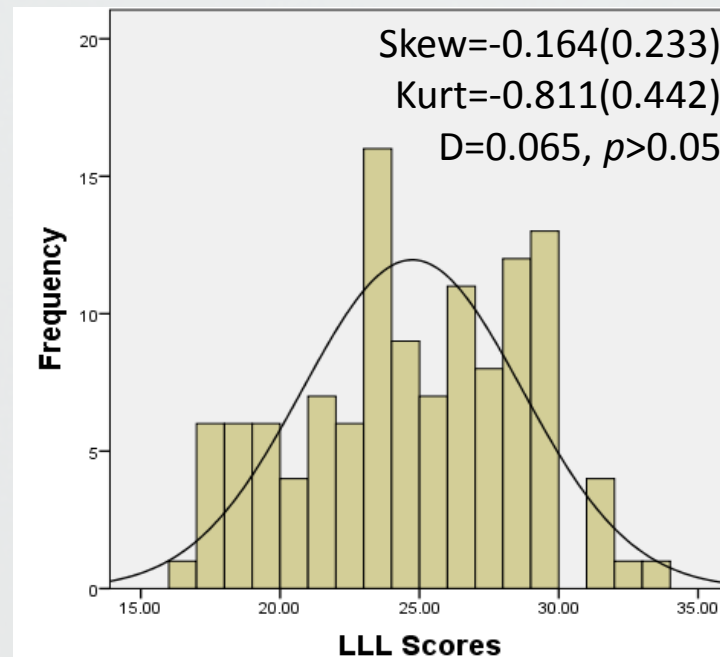
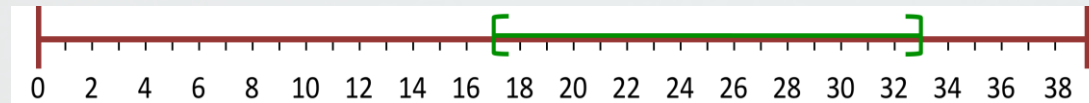
- Encuesta con 27 ítems: Información demográfica, perfil académico, experiencia profesional, experiencia con tecnología, información de contacto

➤ **Aprendizaje a lo largo de la vida (CLLP)**

- Livneh, C. L. (1986). Characteristics Of Lifelong Learners In The Human Service Professions. (Ph.D.), Boston University, Ann Arbor.
- Escala Likert de 36 ítems: Confiabilidad de $\alpha=0.89$ ($\alpha=0.91$ con la muestra original de Livneh)

Test de Normalidad

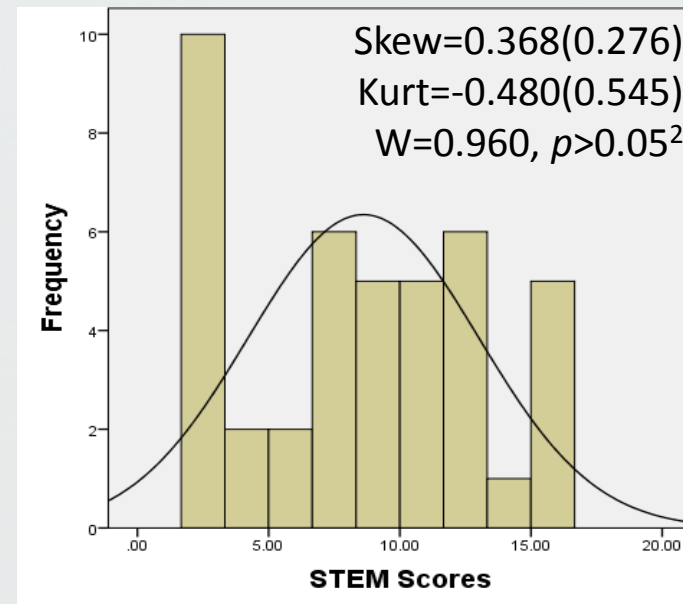
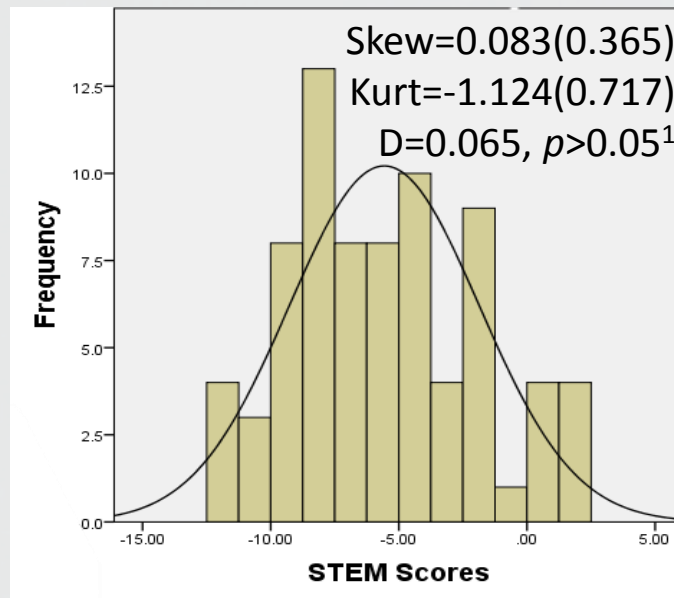
- Aprendizaje a lo largo de la vida
 - Kolmogorov-Smirnov



Test de Normalidad

➤ Career Background

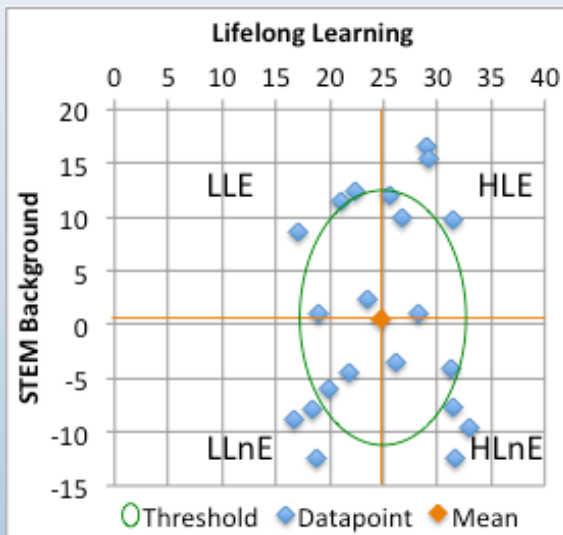
- ¹Kolmogorov-Smirnov
- ²Shapiro-Wilk



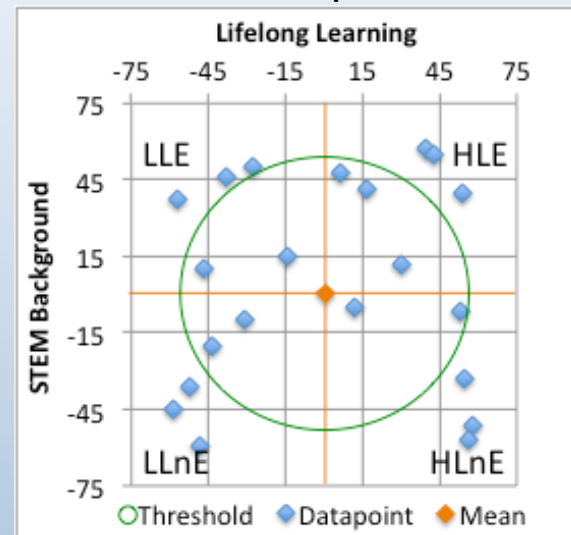
Análisis de Datos

- Identificación de datos atípicos por medio del análisis de nubes $\rightarrow d(o_i, c) > d_{min}$

Datos Crudos



Datos Ranqueados

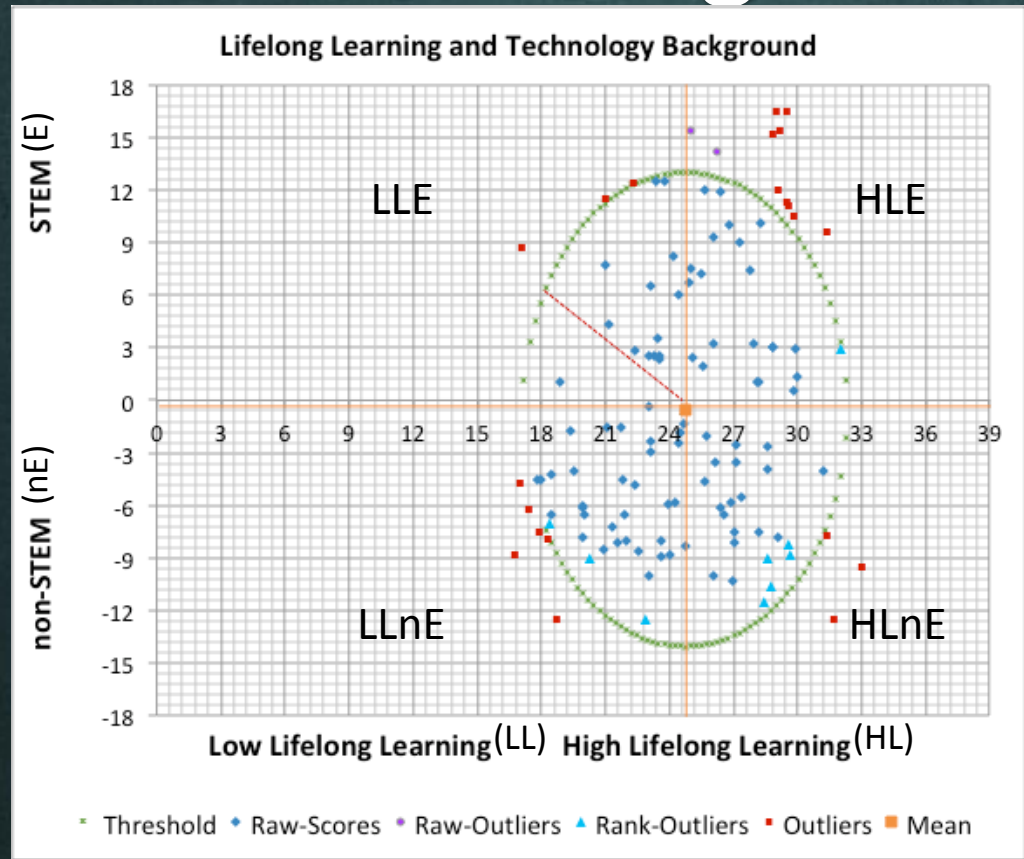


Análisis de Nube – Datos Originales

➤ *Ajuste de d_{min}*

- d_o
- $v_y = 14.50$
- $v_x = 8.14$

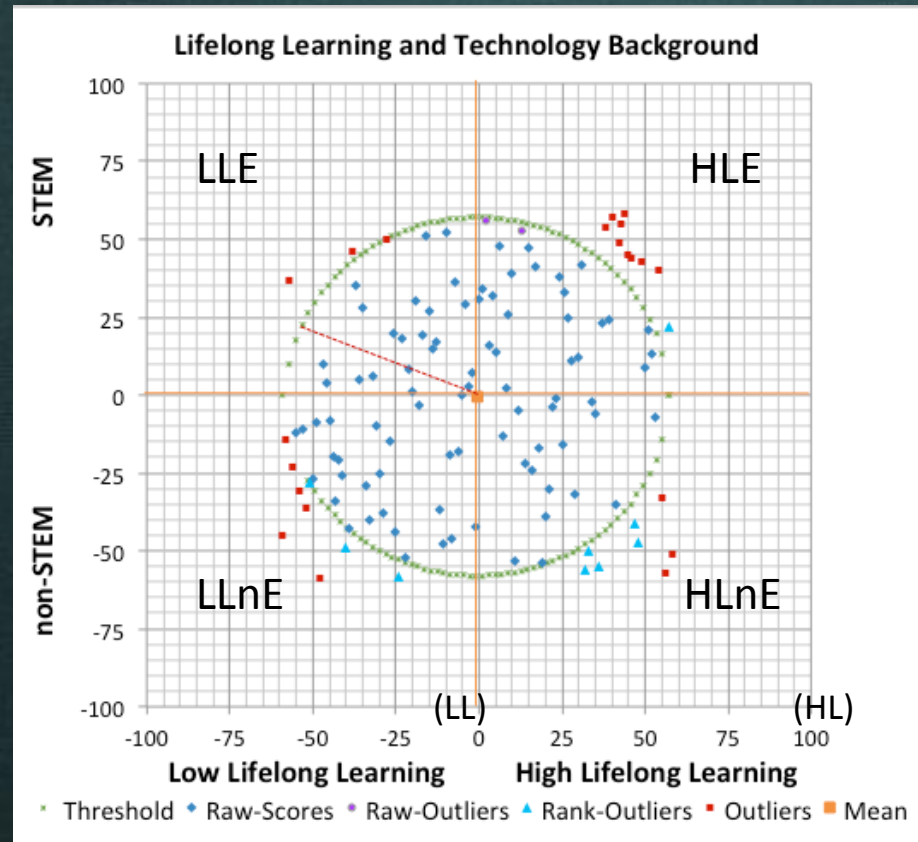
- d_1
- $v_y = 13.54$
- $v_x = 7.60$



Análisis de Nube – Datos Ordenados

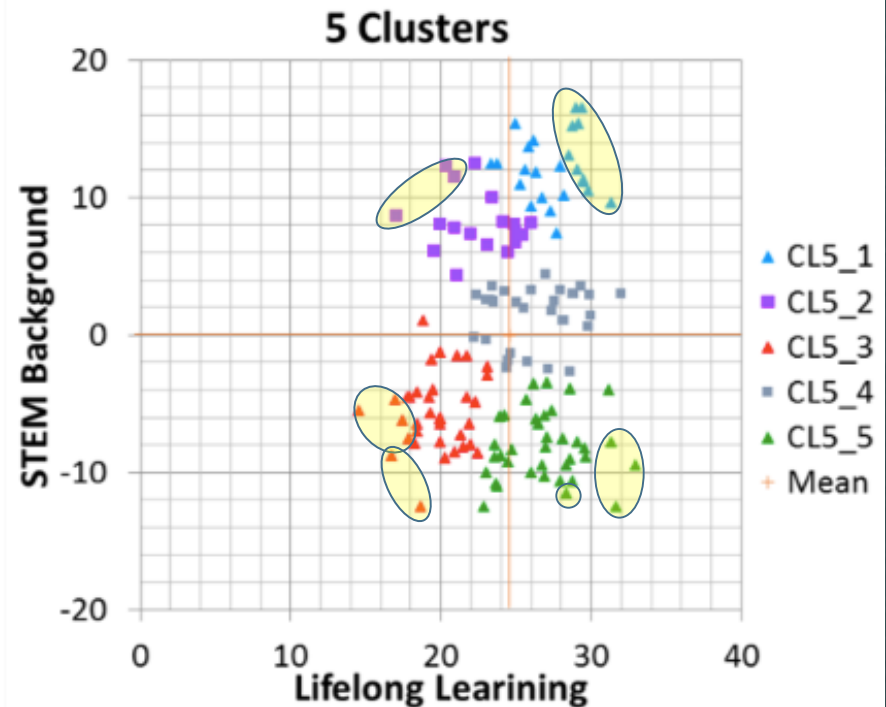
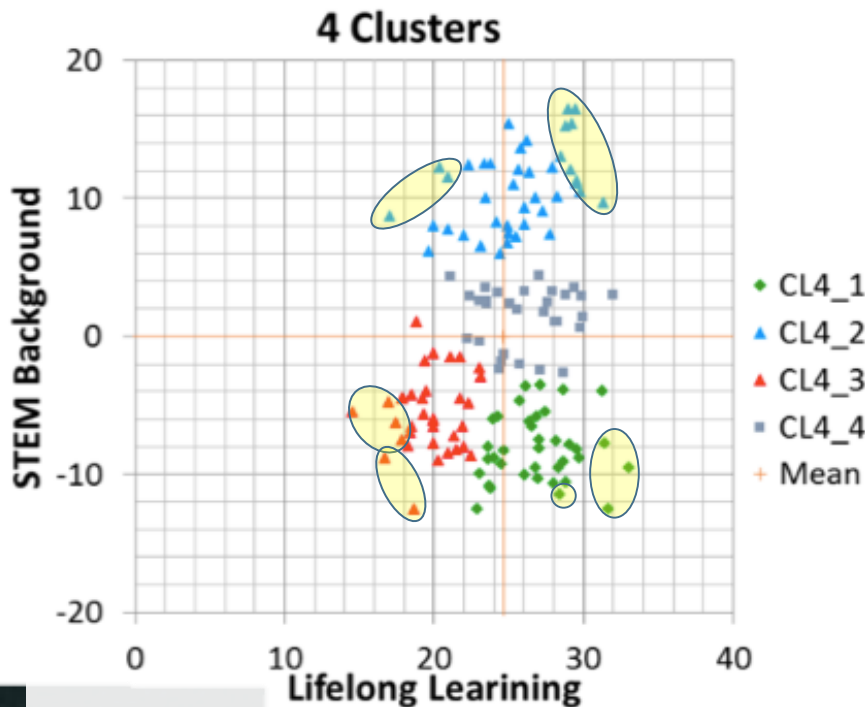
- *Ajuste de d_{min}*
 - d_0
 - $r = 58.5$

- d_1
- $r = 57.5$



Análisis de Grupos

➤ Análisis de nubes usando promedios k



Comparación de Perfiles

	Muestra Inicial	Muestra Intencional
Aprendizaje a lo largo de la vida	HL=27.92, LL=21.34 F (3,114) = 94.24, p < 0.05	HL=30.15, LL=18.49
Perfil Profesional	E=7.47, nE=-6.39 F (3,114) = 104.81, p < 0.05	E=12.56, nE=-8.61
Factores CLLP	HLE=6.07, HLE=5.92, LLE=5.68, LLnE=5.55	HLE=6.38, HLE=6.31, LLE=5.50, LLnE=4.98
• Crecimiento Profesional	Δ LLnE-HL=0.40 F (3,114) = 4.22, p < 0.05	Δ LLnE-HL=1.05
• Causalidad del aprendizaje	Δ LLnE-HLE=0.47 F (3,114) = 4.35, p < 0.05	Δ LLnE-HLE=1.22
• Entorno familiar	Δ LLnE-HL=1.36 HLE F (3,114) = 6.62, p < 0.05	Δ LLnE-HL=2.62

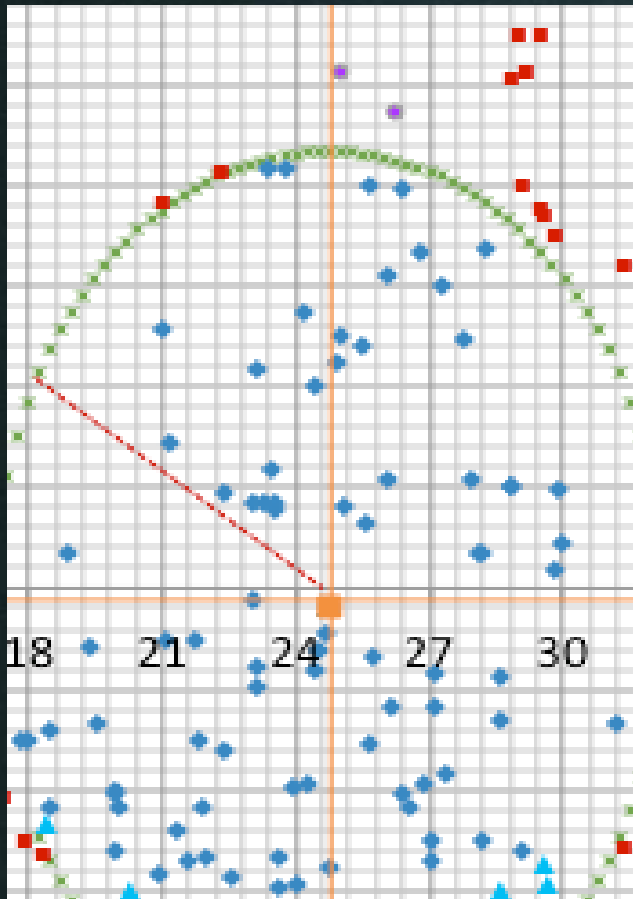
Discusión

- Aplicabilidad del análisis de nubes para identificar datos atípicos
 - En muestras de máxima variación o casos contrastantes
 - Distancia de pertenencia a la nube y forma para selección pueden ser adaptadas según necesidades
 - Útil cuando se tiene más de un criterio de selección
- Implicaciones
 - El poder reducir el sesgo en la selección de muestras intencionales
 - La muestra inicial debe ser al menos $de N=oc/p$ para obtener datos críticos

A tener en cuenta

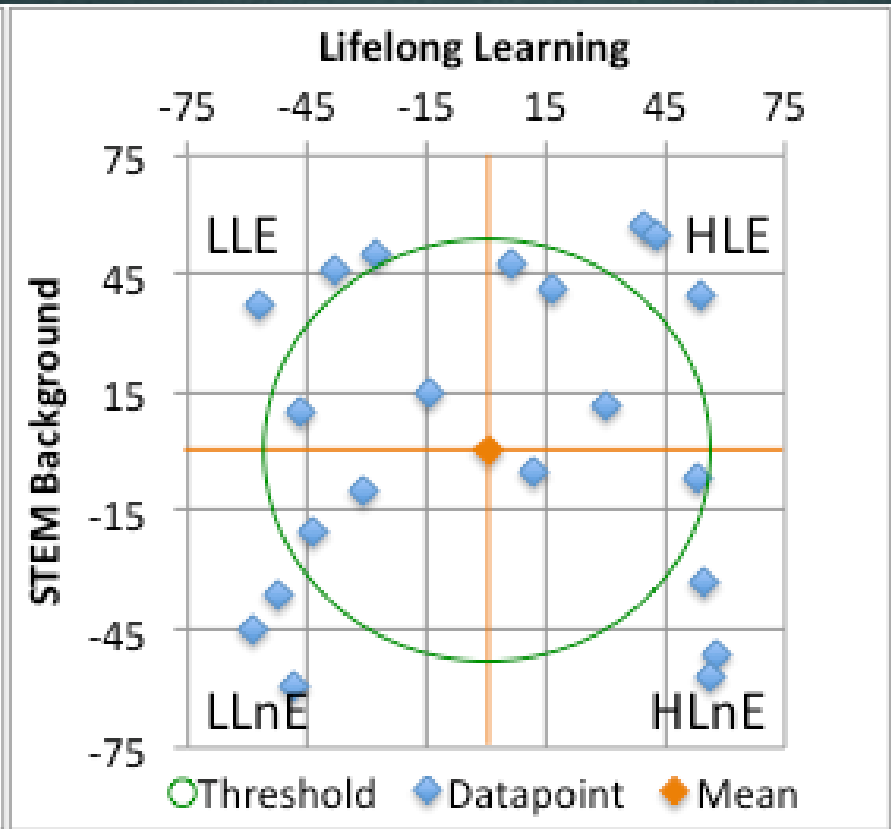
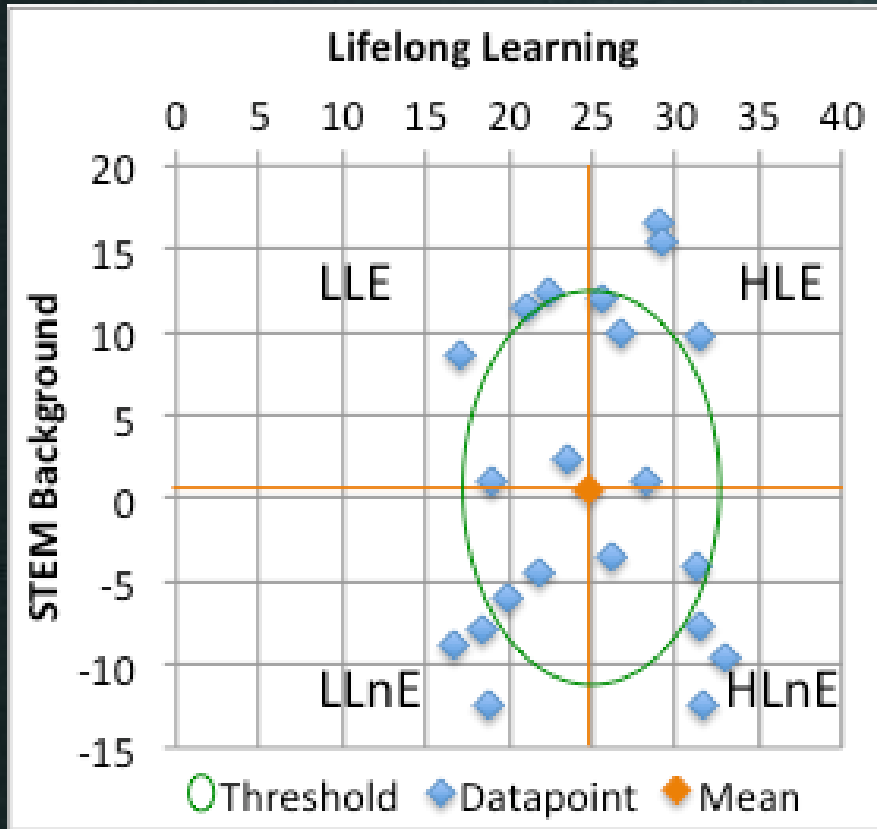
- Limitaciones del método propuesto
 - Falta de control en la selección de grupos o agrupaciones
 - Otro tipos de análisis pueden ser realizados para analizar agrupaciones (Tafur, 2015)
- A Futuro
 - Replicar el estudio para diferentes tamaños de muestra
 - Analizar el método para otro tipo de muestreo intencional (ej. Muestras homogéneas, uso de más variables)
 - Cuantificación de la reducción del sesgo por medio de grupos control y líneas de base

Conclusión



- Realizar un análisis estadístico en la selección de muestras intencionales puede reducir el sesgo inherente en los investigadores cuando seleccionan la muestra de su propio estudio

Gracias!



LL → Aprendizaje a lo largo de la vida Bajo
 HL → Aprendizaje a lo largo de la vida Alto

E → Perfil en Ingeniería o Tecnología
 nE → Perfil en otras disciplinas















Table 3 CLLP Factors for Outlier Profile

Characteristic	LLnE (n=6)	LLE (n=3)	HLnE (n=3)	HLE (n=9)
Professional Growth	<p>↓ 5.60</p> <p>The lowest scored group, more than 1SD below average. Participants agree they believe that keeping updated and competent in their profession is important</p> <p>M= 6.21 SD= 0.59</p>	<p>→ 5.94</p> <p>Slightly below average. Participants agree they believe that keeping updated and competent in their profession is important</p>	<p>↗ 6.61</p> <p>Above average. Participants strongly agree they believe that keeping updated and competent in their profession is important</p>	<p>↗ 6.72</p> <p>The highest scored group, above average. Participants strongly agree they believe that keeping updated and competent in their profession is important</p>
Self-motivated	<p>↘ 5.34</p> <p>The lowest scored group, below average. Participants mildly agree they are motivated and determined to do well</p> <p>M= 5.83 SD= 0.70</p>	<p>↘ 5.40</p> <p>Below average. Participants mildly agree they are motivated and determined to do well</p>	<p>↗ 6.20</p> <p>The highest scored group, above average. Participants agree they are motivated and determined to do well</p>	<p>→ 6.18</p> <p>Slightly above average. Participants agree they are motivated and determined to do well</p>
Educability	<p>↘ 4.91</p> <p>The lowest scored group, below average. Participants mildly agree they have an interest in reading</p> <p>M= 5.72 SD= 0.87</p>	<p>↘ 5.13</p> <p>Below average. Participants mildly agree they have an interest in reading</p>	<p>↑ 6.60</p> <p>The highest scored group, more than 1SD above average. Participants strongly agree they have an interest in reading</p>	<p>→ 5.93</p> <p>Slightly above average. Participants agree they have an interest in reading</p>
Readiness to Change	<p>↘ 5.32</p> <p>The lowest scored group, below average. Participants mildly agree they are able to cope with change</p> <p>M= 5.91 SD= 0.73</p>	<p>→ 5.75</p> <p>Slightly below average. Participants agree they are able to cope with change</p>	<p>↗ 6.33</p> <p>Above average. Participants agree they are able to cope with change</p>	<p>↗ 6.56</p> <p>The highest scored group, above average. Participants strongly agree they are able to cope with change</p>

Tomado de Tafur, 2015, Cap.4








Table 3 (Cont.)

Characteristic	LLnE (n=6)	LLE (n=3)	HLnE (n=3)	HLE (n=9)
Causation for Learning M= 5.53 SD= 0.70	 4.95 The lowest scored group, below average. Participants mildly agree they are involved in learning activities	 5.11 Below average. Participants mildly agree they are involved in learning activities	 6.33 The highest scored group, more than 1SD above average. Participants agree they are involved in learning activities	 6.13 Above average. Participants agree they are involved in learning activities
Familiar Background M= 4.97 SD= 1.67	 3.19 The lowest scored group, more than 1SD below average. Participants mildly disagree their parents participated in learning	 5.22 Slightly above average. Participants mildly agree their parents participated in learning	 6.00 Above average. Participants agree their parents participated in learning	 6.07 The highest scored group, above average. Participants agree their parents participated in learning
Future Orientation M= 6.09 SD= 0.69	 4.64 The lowest scored group, more than 1SD below average. Participants mildly agree they have a desire to advance in their jobs	 5.92 Slightly below average. Participants agree they have a desire to advance in their jobs	 6.58 The highest scored group, above average. Participants strongly agree they have a desire to advance in their jobs	 6.56 Above average. Participants strongly agree they have a desire to advance in their jobs

Outliers in each group scored...

Green number the highest across groups
Red Number the lowest across groups

-  More than 1 SD above population mean
-  Between 0.5 and 1 SD above population mean
-  Less than 0.5 SD from population mean
-  Between 0.5 and 1 SD below population mean
-  More than 1 SD below population mean

Tomado de Tafur, 2015, Cap.4